Regularization in global sound equalization based on effort variation

Stefanakis, Nick; Sarris, John; Jacobsen, Finn

Published in:
Acoustical Society of America. Journal

Link to article, DOI:
10.1121/1.3158926

Publication date:
2009

Document Version
Publisher's PDF, also known as Version of record

Regularization in global sound equalization based on effort variation

Nick Stefanakis[a] and John Sarris
School of Electrical and Computer Engineering, National Technical University of Athens, Heroon Polytechniou 9, 157 73 Athens, Greece

Finn Jacobsen
Department of Electrical Engineering, Acoustic Technology, Technical University of Denmark, Ørsteds Plads, Building 352, DK-2800 Kongens Lyngby, Denmark

(Received 20 June 2008; revised 14 May 2009; accepted 28 May 2009)

Sound equalization in closed spaces can be significantly improved by generating propagating waves that are naturally associated with the geometry, as, for example, plane waves in rectangular enclosures. This paper presents a control approach termed effort variation regularization based on this idea. Effort variation equalization involves modifying the conventional cost function in sound equalization, which is based on minimizing least-squares reproduction errors, by adding a term that is proportional to the squared deviations between complex source strengths, calculated independently for the sources at each of the two walls perpendicular to the direction of propagation. Simulation results in a two-dimensional room of irregular shape and in a rectangular room with sources randomly distributed on two opposite walls demonstrate that the proposed technique leads to smaller global reproduction errors and better equalization performance at listening positions outside of the control region compared to effort regularization and compared to a simple technique that involves driving groups of sources identically.
 [DOI: 10.1121/1.3158926]

PACS number(s): 43.55.Br, 43.38.Md, 43.60.Pt [NX] Pages: 666–675

I. INTRODUCTION

The purpose of equalization in room acoustics is to compensate for the undesired modification that an enclosure introduces to signals as, for example, audio or speech. Traditional multi-channel methods introduce digital filters to preprocess the input signal before it is fed to a set of loudspeakers so that the spectral coloration and the reverberation tail associated with the transmission path are reduced.[1,2] Generally, equalization is focused on improving two different attributes of the listening response: the reverberation time and the magnitude response. Modal equalization attempts to control the modal decay of low-frequency modes so that they correspond to a target reverberation time.[3,4] This method is based on a rearrangement of the poles of the listening response. It is a spatially robust method in the sense that increasing the decay rate of a modal resonance at a single position in the room results in an increment of the decay rate at other positions as well.[5] Magnitude response equalization attempts to reduce the unevenness associated with the peaks and dips in the spectrum of the listening response. Generally, the process aims to design the source input signal so that the obtainable signals at a set of receiving positions approximate a set of desired signals.[6,7]

Based on the latter approach, it has been shown that equalization in a large region of a room can be favored by the reproduction of a freely propagating plane wave.[8–13] This is achieved by appropriate positioning of sound sources at two opposite walls perpendicular to the direction of propagation. In this way the equalization can be extended to a spatial region that covers almost the complete volume of the room. It has been shown theoretically that for the successful generation of a plane wave in a rectangular room, the sources at each side are in phase and their amplitudes assume well-defined ratios.[10] These ratios reflect the natural effect of the nearby reflecting surfaces, varying according to the source position inside the room. Also, for the case of a perfectly rectangular room with uniform acoustic wall admittance, optimum source locations as well as favorable listening planes have been identified.[13] In the same work it has been shown that by exploiting the symmetries in the rectangular room, an ideal source distribution can be found. In that case, a single equalization filter at each side and a limited control sensor array are enough to achieve a successful equalization over an area covering almost the complete volume of the room. These observations suggest that the source filters at each one of the two "playing" walls are highly correlated and this correlation is desired in order to achieve a spatially extended equalization.

This work presents a modification to the cost function used in the traditional least-squares approximation. The proposed technique, termed *effort variation regularization*, is based on the general form of Tikhonov regularization[14] and works by penalizing the squared deviations between source strengths.

In the frequency domain, penalization of the complex source strength variations can be realized using square tridi-

[a]Author to whom correspondence should be addressed. Electronic mail: nstefan@mobile.ntua.gr

agonal or bidiagonal differential operators instead of the unitary matrix used in traditional effort regularization.[14] Simulation results for a room of irregular shape demonstrate the applicability of the technique when the conditions for equalization are more difficult than in perfectly rectangular rooms.

Following the traditional least-squares approach for the equalization of broadband signals, effort variation regularization can easily be applied in the time domain using finite impulse response (FIR) equalization filters. Simulation results for a three-dimensional rectangular room show that good global equalization can be achieved with a few control sensors placed outside the listening area, without restraining the presence and motion of listeners inside the room.

II. CONTROL MODEL

Suppose that it is desired to equalize the sound field in a spatial region in an enclosure with L reproduction sources. The sound pressure in the region is sampled by M monitor sensors placed at $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_M\}$, and this provides a measure of the performance of reproduction in the entire listening space. The pressure at the monitoring sensors subject to the L source excitations can be written as[7]

$$\mathbf{p}_M = \mathbf{Z}_M \mathbf{q}, \tag{1}$$

where $\mathbf{p}_M$ is a column vector with the M complex sound pressures at the monitor sensors (Pa), $\mathbf{q}$ is a column vector with the complex strengths of the L sources (m^3/s), and $\mathbf{Z}_M$ is an $M \times L$ transfer matrix with the transfer functions between the L sources and the M monitor sensors. A small group of N control sensors is selected from the M monitor sensors at $\{\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N\}$, covering a small region centered inside the listening area. It is assumed that this compact control sensor array represents a more practical sound reproduction system that occupies less space and requires less equipment and input channels. This system is informed about the performance of reproduction in the controlled region by the difference between the desired sound pressure and the actual reproduced sound pressure at the control sensors as expressed by

$$\mathbf{e} = \mathbf{p}_d - \mathbf{Z}\mathbf{q}, \tag{2}$$

where $\mathbf{p}_d$ is the vector with the desired sound pressures at the N control sensors, and $\mathbf{Z}$ is the transfer matrix with the transfer functions between the L sources and the N control sensors.

A. Regularization techniques

Equalization is related to the inverse problem of reconstructing the strengths of a number of sources given the transfer matrix and the desired responses at a number of receiving positions in the room. When the number of sensors is less than the number of sources the linear problem is underdetermined and reconstruction of the source strengths requires the use of a regularization technique. One of the most common techniques used in active control is effort regularization, which is based on the standard form of Tikhonov regularization[14–16] and the cost function

$$J^{(\mu)} = (\mathbf{Z}\mathbf{q} - \mathbf{p}_d)^H (\mathbf{Z}\mathbf{q} - \mathbf{p}_d) + \mu \mathbf{q}^H \mathbf{q}, \tag{3}$$

where μ is a positive scalar that weights the penalty term in the cost function. Minimization of this cost function gives a solution to the underdetermined system. Furthermore, it has been shown that a proper choice of the regularization parameter μ has the ability to enlarge the effective area of equalization inside a room.[12] The optimum source strength vector for this regularization technique is

$$\mathbf{q}_o^{(\mu)} = (\mu \mathbf{I} + \mathbf{Z}^H \mathbf{Z})^{-1} \mathbf{Z}^H \mathbf{p}_d. \tag{4}$$

In search of more efficient penalization techniques, power output regularization has recently been presented and associated with the general form of Tikhonov regularization.[12] Used in a similar equalization task, penalization of the total power output favored the reproduction of the plane propagating wave in the room, resulting in an increment of the effective equalization area.

Based also on the general form of Tikhonov regularization, effort variation regularization relies on the use of the cost function

$$J^{(h)} = (\mathbf{Z}\mathbf{q} - \mathbf{p}_d)^H (\mathbf{Z}\mathbf{q} - \mathbf{p}_d) + h(\mathbf{D}\mathbf{q})^H (\mathbf{D}\mathbf{q}), \tag{5}$$

where h is a positive scalar that weights the effort variation penalty term in the cost function, and the weight matrix $\mathbf{D}$ represents a discrete first or second order differential operator. Such matrices are usually of the form[17]

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & 0 & \ldots \\ 0 & -1 & 1 & 0 & \ldots \\ 0 & 0 & -1 & 1 & \ldots \\ 0 & 0 & 0 & -1 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{6}$$

and

$$\mathbf{D} = \begin{bmatrix} -1 & 1 & 0 & 0 & \ldots \\ 1 & -2 & 1 & 0 & \ldots \\ 0 & 1 & -2 & 1 & \ldots \\ 0 & 0 & 1 & -2 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \tag{7}$$

for first and second order variation penalties, respectively. If the quantity $(\mathbf{D}\mathbf{q})^H(\mathbf{D}\mathbf{q})$ is expanded in terms of the elements of the vector $\mathbf{q}$, it can be seen that such matrices penalize variations between the strengths of adjacent sources and thus force the system to a more uniform solution, which in the extreme case would lead to all the sources having equal strengths ($q_1 = q_2 = \cdots = q_L$). Under the condition that the matrix $h\mathbf{D}^H\mathbf{D} + \mathbf{Z}^H\mathbf{Z}$ is invertible, the optimum solution can be derived as

$$\mathbf{q}_o^{(h)} = (h\mathbf{D}^H\mathbf{D} + \mathbf{Z}^H\mathbf{Z})^{-1} \mathbf{Z}^H \mathbf{p}_d. \tag{8}$$

It should be noted that for such differential operators the matrix $\mathbf{D}^H\mathbf{D}$ is singular, and therefore too strong penalization can be expected to harm the calculation of $(h\mathbf{D}^H\mathbf{D} + \mathbf{Z}^H\mathbf{Z})^{-1}$. Also, in contrast to the case of effort regularization, the numbering of the sources and their positioning in the source strength vector $\mathbf{q}$ is of great importance, and, as will be

shown, it must be consistent with the source array geometry.

B. Reproduction error criteria

The achieved quality of the equalization is measured in the entire listening space with the use of the M monitor sensors. The global reproduction error is calculated as the square root of the mean square value of the errors at the monitor sensors normalized by the energy of the desired sound field at the monitor sensors $\mathbf{p}_{d,M}$,

$$\mathrm{ELS}_m^{(j)} = \left(\frac{\mathbf{e}_M^H \mathbf{e}_M}{\mathbf{p}_{d,M}^H \mathbf{p}_{d,M}} \right)^{1/2} = \left(\frac{(\mathbf{p}_{d,M} - \mathbf{Z}_M \mathbf{q}_o^{(j)})^H (\mathbf{p}_{d,M} - \mathbf{Z}_M \mathbf{q}_o^{(j)})}{\mathbf{p}_{d,M}^H \mathbf{p}_{d,M}} \right)^{1/2}. \quad (9)$$

Here $\mathbf{e}_M = \mathbf{p}_{d,M} - \mathbf{Z}_M \mathbf{q}_o^{(j)}$ and $j = \mu$ or h, which implies effort regularization and effort variation regularization, respectively. For the given definition of the reproduction error, and assuming that the equalization is achieved by the generation of a propagating plane wave, a value up to 0.3 implies that the deviations between the reproduced sound pressure and the desired pressure are within an interval of ± 3 dB, whereas a value up to 0.5 implies deviations within ± 6 dB.

The performance of the reproduction is also measured at the N control sensors as

$$\mathrm{ELS}_c^{(j)} = \left(\frac{(\mathbf{p}_d - \mathbf{Z}\mathbf{q}_o^{(j)})^H (\mathbf{p}_d - \mathbf{Z}\mathbf{q}_o^{(j)})}{\mathbf{p}_d^H \mathbf{p}_d} \right)^{1/2}. \quad (10)$$

In addition to the control approaches described above it is also interesting to define an *ideal system* that uses the information from all M monitor sensors in the adaptation of the optimum source strengths by minimizing the global cost function

$$J^{(\mathrm{id})} = \mathbf{e}_M^H \mathbf{e}_M. \quad (11)$$

The optimum source strengths for this control approach are given by

$$\mathbf{q}_o^{(\mathrm{id})} = (\mathbf{Z}_M^H \mathbf{Z}_M)^{-1} \mathbf{Z}_M^H \mathbf{p}_{d,M}. \quad (12)$$

This corresponds to the best that can be achieved in the ideal, unrealistic case where control sensors covering the entire listening space are used. No kind of regularization is included in the calculation of the optimum source strengths given by Eq. (12). Under the examined conditions, the matrix $\mathbf{Z}_M$ is always overdetermined and $\mathbf{Z}_M^H \mathbf{Z}_M$ is positive definite. The global reproduction performance of the ideal system can also be judged by the global reproduction error by substituting $\mathbf{q}_o^{(j)}$ with $\mathbf{q}_o^{(\mathrm{id})}$ in Eq. (9).

III. SIMULATIONS IN A TWO-DIMENSIONAL ROOM

A. Conditions for the simulations

The purpose of this section is to investigate equalization in a non-rectangular room, which is modeled using the boundary element method.[18,19] A sum of 3784 linear triangular elements is used in order to mesh the physical boundary of the room, shown in Fig. 1. The vertical dimension of the room is kept relatively small ($L_z = 0.1$ m) so that the equalization problem is actually reduced to a two-dimensional one. A constant acoustic wall impedance of $120\rho_0 c$, where $\rho_0 = 1.204$ kg/m^3 is the density of air and $c = 344$ m/s is the speed of sound, has been assumed. Eleven sound sources with monopole characteristics are used to control the sound field; five point sources are placed on the left wall at $x_L = 0.1$ m, and the remaining six sources are placed near the right wall at $x_R = 3.4$ m. The equalization is examined in the listening area, which is defined as a rectangle with the lower left corner at (0.7, 0.3, 0.1) m and the right upper corner at (2.8, 1.8, 0.1) m. A sum of 352 monitor sensors is spread inside the listening area in order to monitor the sound pressure, and a square grid of nine control sensors is used near the middle of the room in order to optimize the complex source strengths. The central control sensor is placed at (2, 1, 0.1) m. The distance between the monitor sensors as well as between the control sensors is 0.1 m in both the x- and the y-direction. In this configuration the control region covers less than 4% of the listening area. The purpose of the arrangement is to reproduce a plane wave that travels in the x-direction, simulating a sound field of constant amplitude that varies as e^{-jkx}.

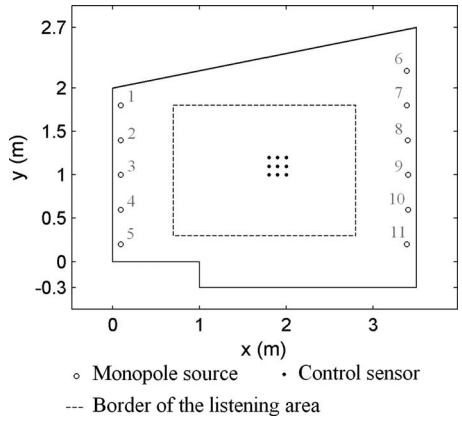

FIG. 1. Configuration of the equalization system in a two-dimensional non-rectangular room. The lower left monopole source coordinates are (0.1, 0.2, 0.1), and the upper left are (0.1, 1.8, 0.1), while the lower right and upper right source coordinates are (3.4, 0.2 0.1) and (3.4, 2.6, 0.1), respectively. The source spacing is 0.4 m at both sides.

B. The ideal system

Reproduction of the plane wave subject to minimization of the reproduction error at the M monitor sensors is accomplished successfully. Implementation of the ideal system is based on knowledge of the transfer functions from all sources to all monitor sensors, something that would be impractical in a real situation. However, this approach is worth examining because it provides useful observations about the correlation in the optimum complex source strengths required for the generation of the plane wave. The optimum source amplitudes are shown in Figs. 2(a) and 2(c) for the five left and the six right point sources, respectively, and the optimum phases are shown in Figs. 2(b) and 2(d). Some deviations between the source amplitudes can be seen in Figs. 2(a) and 2(c). The deviations are seen to be smaller at the left wall than at the right one. On the other hand, inspec-

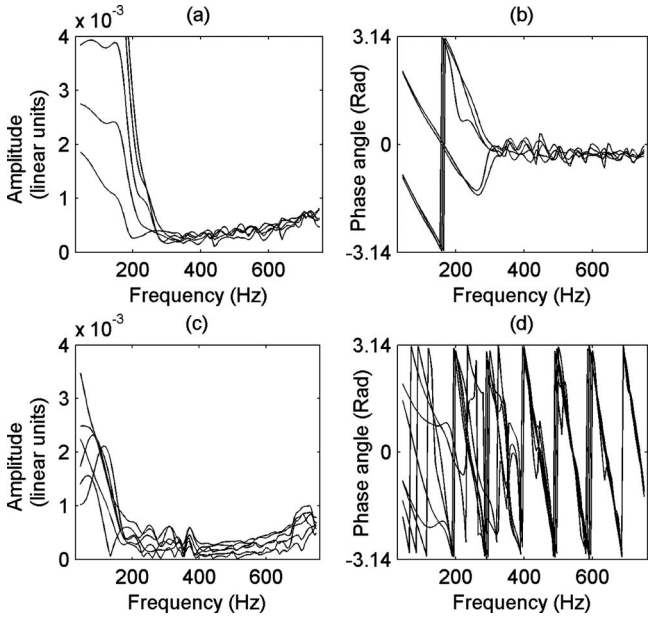

FIG. 2. Amplitudes (a) and phases (b) of the left sources (at $x=0.1$ m), and amplitudes (c) and phases (d) of the right sources (at $y=3.4$ m) with the ideal system.

tion of Figs. 2(b) and 2(d) shows that the sources at each side tend to have identical phases, in particular, above 300 Hz. Note that the phases at the left side are stabilized around zero whereas the phases at the right side tend to vary linearly with the frequency according to $-e^{-jkx_R}$. This means that the sources are 180° out of phase with the desired sound pressure at $x=x_R$, which is necessary to avoid reflections from the receiving wall. This behavior, which has already been observed and explained theoretically by Santillán[8] for the case of rectangular rooms, remains a desired condition for good global equalization also in irregularly shaped rooms.

C. Implementation of effort variation regularization

For the given arrangement and the source numbering shown in Fig. 1, the effort variation penalty term of Eq. (5) is implemented with the matrix

$$\mathbf{D} = \begin{bmatrix} \mathbf{D}_L & \mathbf{0}_{6\times 6} \\ \mathbf{0}_{5\times 5} & \mathbf{D}_R \end{bmatrix}, \tag{13}$$

where

$$\mathbf{D}_L = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \tag{14}$$

$$\mathbf{D}_R = \begin{bmatrix} -1 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}, \tag{15}$$

and $\mathbf{0}_{5\times 5}$ and $\mathbf{0}_{6\times 6}$ are zero matrices. These second order differential operators penalize the variation between the strengths of adjacent sources. Note that with the numbering shown in Fig. 1 sources 1 and 5 have only one neighbor and are therefore linked only to sources 2 and 4, respectively, whereas source 2, for example, is linked to both sources 1 and 3. It is also interesting to observe from Eq. (13) that none of the sources 1–5 is linked to any of the opposite sources 6–11. This suggests that the deviations in the solution are penalized independently at each side.

D. Equalization with two coupled source arrays

Inspired by previous work on sound equalization in a rectangular room[13] the idea of coupling the sources at each side of the room is also examined. In this particular case this means that the five left sources are driven with the same amplitude and phase, and the same holds for the six right sources. This technique is evidently similar to using two independent "column" loudspeakers,[20] one at each side of the room. In the frequency domain, only two complex source strengths, one for the left source array and one for the right one, should be estimated. This corresponds to solving an overdetermined system

$$\mathbf{q}_{\text{coupled}} = \begin{bmatrix} q_L \\ q_R \end{bmatrix} = (\mathbf{Z}^H_{\text{coupled}}\mathbf{Z}_{\text{coupled}})^{-1}\mathbf{Z}^H_{\text{coupled}}\mathbf{p}_d, \tag{16}$$

where q_L and q_R are the left and right coupled source array strengths, respectively, and $\mathbf{Z}_{\text{coupled}}$ is the 9×2 matrix carrying the acoustic transfer functions from each source array to the control sensors inside the room. This approach thus assumes a-priori known conditions about the source strengths, giving results that would be similar to those derived by effort variation regularization in the extreme case where $h\rightarrow\infty$.

E. Reproduction performance

The system is now tested for global equalization subject to minimization of reproduction errors only at the nine control sensors. Examination of the reproduction error at the control sensor locations shows that the increment of the regularization parameters μ and h is followed by reduction in the quality of equalization at these positions. Nevertheless, the accuracy of the reproduction in the entire listening area is improved for a non-trivial value of the regularization parameters with both the proposed and the conventional technique. This can be seen in Fig. 3, where the global reproduction error is plotted as a function of the regularization parameter for effort regularization and effort variation regularization at 520 Hz. The values of the regularization parameters in Eqs. (4) and (8) were set equal to $\mu=450$ and $h=900$, and the global reproduction error was determined between 30 and

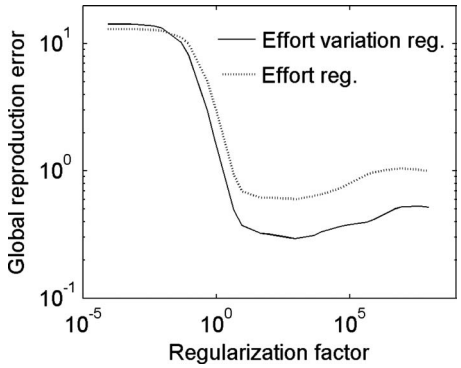

FIG. 3. Global reproduction error as a function of the regularization parameter for effort regularization and effort variation regularization for second order differential operator at 520 Hz.

750 Hz. Figure 4(a) shows the global reproduction error also for the ideal and for the coupled source system, and Fig. 4(b) shows the reproduction error at the control sensor locations for both regularization techniques as well as for the coupled source system. Both regularization techniques achieve an almost perfect equalization result in the control region in the entire frequency range. However, whereas the reproduction error in the control region is of the same order of size both with effort regularization and effort variation regularization, the global performance of these two techniques is evidently different. It can be seen in Fig. 4(a) that effort regularization leads to large global reproduction errors above 300 Hz. This is translated into a rapid decline of the reproduction performance when moving away from the control positions. Whereas this deterioration increases with the frequency with the traditional regularization technique it can be seen that the proposed regularization technique achieves much smaller re-

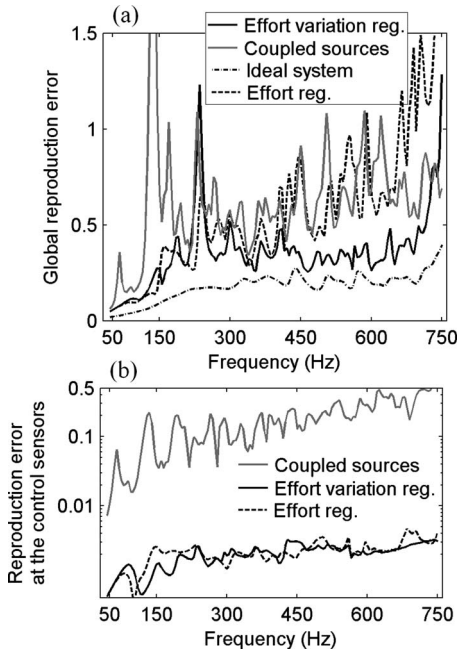

FIG. 4. (a) Global reproduction error as a function of the frequency for the ideal system, effort regularization, effort variation regularization, and coupled source arrays. (b) Reproduction error at the control sensors for effort regularization, effort variation regularization, and coupled source array.

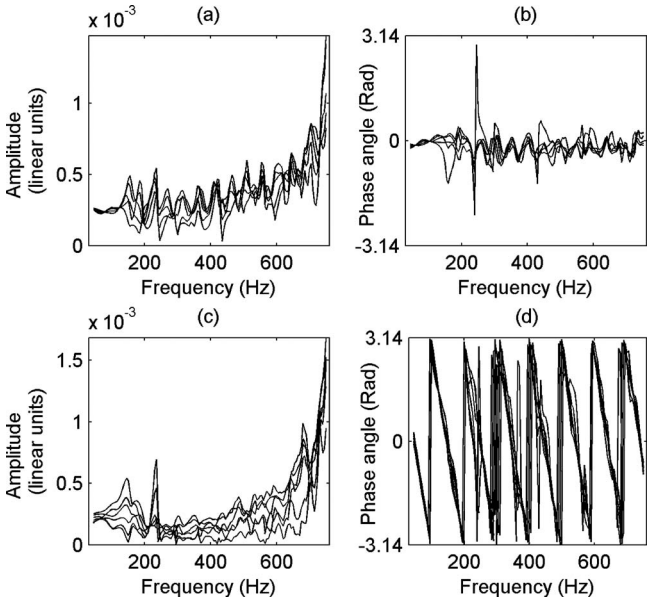

FIG. 5. Amplitudes (a) and phases (b) of the left sources, and amplitudes (c) and phases (d) of the right sources with effort variation regularization.

production errors from 300 to 700 Hz. This indicates that the proposed control approach gives better equalization results at the monitor positions outside the control region. This demonstrates an important advantage in this frequency region, but it should be mentioned that all techniques apart from the ideal system exhibit a poor global performance around 230 Hz. The frequency of this error peak appears to be depending on the location of the control sensors. In any case it was observed that proper adjustment of the regularization factors μ and h reduces the global reproduction error below 0.5 around 230 Hz both for effort regularization and effort variation regularization. In the simulation results presented above, a constant value of the regularization factor was used in the entire frequency range.

Examination of the coupled source system shows that it is unable to provide a good global equalization result, although some improvement is observed at the highest frequencies of the investigation. This indicates that the success of the proposed technique compared to the case of the coupled source array is due to imposing a desired correlation without necessarily preventing deviations that are required for the adaptation to the particulars of the room. The individual effect of each regularization technique in each derived solution $\mathbf{q}_o^{(h)}$ and $\mathbf{q}_o^{(\mu)}$ can be seen in Figs. 5 and 6. The variation in the source amplitudes at the left and the right side can be seen in (a) and (b), and the variation in the source phases is shown in (c) and (d). Comparing the source phases shown in Figs. 5 and 6 shows that the effort variation penalty term has reduced the phase deviations at each side. However, whereas the amplitude deviations at the left side are reduced compared to effort regularization, an evident deviation is observed at the right source amplitudes. In agreement with previous observations this shows that effort variation has acted by equalizing the source phases without preventing the amplitude deviation that is necessary for room compensation.

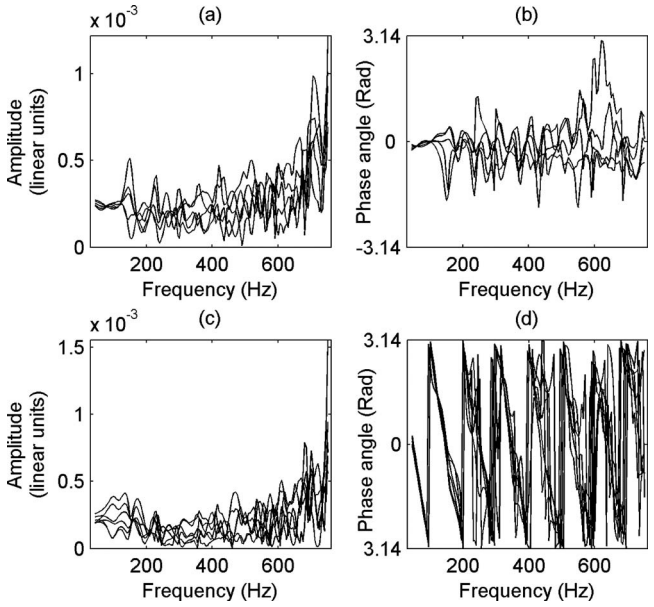

FIG. 6. Amplitudes (a) and phases (b) of the left sources, and amplitudes (c) and phases (d) of the right sources with effort regularization.

IV. EFFORT VARIATION REGULARIZATION IN TIME DOMAIN EQUALIZATION OF BROADBAND SIGNALS

The following is intended to demonstrate the use of effort variation regularization in the time domain and also involves applying the proposed technique to cases of more complicated source distributions, such as those required for three-dimensional rooms. To overcome the time and memory restrictions imposed by the boundary element method, the acoustic transfer functions are calculated here on the basis of an analytical model which is appropriate for rectangular rooms.

A. Time domain implementation of equalization

The following analysis is based on the matrix formulation by Elliott and Nelson.[6] A block diagram of the equalization system is shown in Fig. 7. The aim is to design L digital FIR control filters, one for each sound source, with impulse responses $h_l(n)$ such that the obtainable signal $\hat{d}_m(n)$ at the mth control sensor $(1 \le m \le N)$ is the best approximation to the desired signal $d_m(n)$ at the sensor. (Although the index m has so far been associated with the monitor sensors, it is used here to avoid confusion with the discrete time index n.) Here $d_m(n)$ is a delayed version (by δ_m samples) of the original

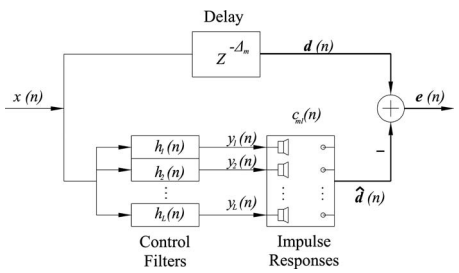

FIG. 7. Block diagram for equalization of broadband signals.

input signal $x(n)$. Since a plane wave traveling in the y-direction is desired, the duration of the delay depends on the ordinate of the listening position corresponding to the propagation of the plane wave.

The impulse response between the input of the lth source and the output of the mth sensor is modeled by an FIR filter with J coefficients, represented by $c_{ml}(n)$ in Fig. 7. The signal $\hat{d}_m(n)$ detected by the mth sensor can be expressed as

$$\hat{d}_m(n) = \sum_{l=1}^{L} \sum_{i=1}^{I} h_l(i) r_{ml}(n-i), \tag{17}$$

where $h_l(i)$ is the ith coefficient of the FIR control filter whose output is the input to the lth loudspeaker, and $r_{ml}(n)$ is the reference signal produced by convolving the input signal $x(n)$ with the impulse response $c_{ml}(n)$. Each control filter is assumed to have an impulse response of I samples. In vector notation Eq. (17) can be written as

$$\hat{d}_m(n) = \mathbf{h}^T(0)\mathbf{r}_m(n) + \mathbf{h}^T(1)\mathbf{r}_m(n-1) + \cdots + \mathbf{h}^T(I-1)\mathbf{r}_m(n-I+1), \tag{18}$$

where a composite tap weight vector and a reference signal vector have been defined by

$$\mathbf{h}^T(i) = \begin{bmatrix} h_1(i) & h_2(i) & \cdots & h_L(i) \end{bmatrix} \tag{19}$$

and

$$\mathbf{r}_m^T(n) = \begin{bmatrix} r_{m1}(n) & r_{m2}(n) & \cdots & r_{mL}(n) \end{bmatrix}. \tag{20}$$

The final vector that contains all the I coefficients of all the L control FIR filters can be written as

$$\mathbf{a}^T = \begin{bmatrix} \mathbf{h}^T(0) & \mathbf{h}^T(1) & \cdots & \mathbf{h}^T(I-1) \end{bmatrix}, \tag{21}$$

and the error vector signal at the N control sensors can be expressed as

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{R}(n)\mathbf{a}, \tag{22}$$

where

$$\mathbf{R}(n) = \begin{bmatrix} \mathbf{r}_1^T(n) & \mathbf{r}_1^T(n-1) & \cdots & \mathbf{r}_1^T(n-I+1) \\ \mathbf{r}_2^T(n) & \mathbf{r}_2^T(n-1) & \cdots & \mathbf{r}_2^T(n-I+1) \\ & & \vdots & \\ \mathbf{r}_M^T(n) & \mathbf{r}_M^T(n-1) & & \mathbf{r}_M^T(n-I+1) \end{bmatrix} \tag{23}$$

is the matrix of filtered reference signals. The optimal coefficients of the control filters are determined by minimizing a performance index defined by[21]

$$J = E\{\mathbf{e}^T(n)\mathbf{e}(n) + \lambda \mathbf{a}^T(n)\mathbf{W}\mathbf{a}(n)\}, \tag{24}$$

where E is the expectation operator, $\mathbf{W}$ is a generally symmetric matrix that defines the type of penalization in the cost function, and λ is the regularization parameter that defines the weighting assign to the penalty term. It can be seen that the performance index can be expressed as a quadratic function in terms of all the individual coefficients in the equalization filters. This performance index has a unique global minimum that corresponds to the optimal control filter coefficients

$$\mathbf{a}_o = (E\{\mathbf{R}(n)^T\mathbf{R}(n) + \lambda\mathbf{W}\})^{-1}E\{\mathbf{R}(n)^T\mathbf{d}(n)\}. \quad (25)$$

It should be noticed that Eq. (25) gives the optimal solution in a statistical sense. The only required information about the source input signal is the autocorrelation function. In the case when $\mathbf{W}=\mathbf{I}$ (the identity matrix), the second term in the expression for the cost function, Eq. (24), becomes identical to the effort penalty term that has been used in similar active control applications.[14,15,21] The optimum control filters in this case are specified in terms of the value of the regularization parameter μ as

$$\mathbf{a}_o^{(\mu)} = (E\{\mathbf{R}(n)^T\mathbf{R}(n) + \mu\mathbf{I}\})^{-1}E\{\mathbf{R}(n)^T\mathbf{d}(n)\}. \quad (26)$$

The transformation of the proposed technique to the form necessary for the time domain equalization of random signals is easily accomplished with proper weighting of the filter coefficients. Again the differential operators must be constructed so as to be in agreement with the source numbering and the architecture of the filter vector $\mathbf{a}$. Considering that the sources are numbered continuously, as in the example presented previously in the frequency domain, the matrix $\mathbf{W}$ can be straightforwardly composed by similar differential operators $\mathbf{D}$ as those defined in Eqs. (14) and (15),

$$\mathbf{W} = \mathbf{\Delta}^T\mathbf{\Delta}, \quad (27)$$

where $\mathbf{\Delta}$ is square matrix of order $I \cdot L$ defined as

$$\mathbf{\Delta} = \begin{bmatrix} \mathbf{D} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{D} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{D} \end{bmatrix}. \quad (28)$$

By inspecting the structure of $\mathbf{\Delta}$ and the architecture of the composite filter vector $\mathbf{a}$ in Eqs. (19) and (21) it can be seen that the penalty term in the cost function in Eq. (24) can be written as

$$\mathbf{a}^T\mathbf{\Delta}^T\mathbf{\Delta}\mathbf{a} = \sum_{i=0}^{I-1} \mathbf{h}^T(i)\mathbf{D}^T\mathbf{D}\mathbf{h}^T(i). \quad (29)$$

This implies that the symmetrical matrix $\mathbf{W}$ used in the time domain measures the deviation between adjacent source filter coefficients in the same way as the matrix $\mathbf{D}^T\mathbf{D}$ measures the deviation between adjacent source strengths in the frequency domain. Under the condition that the matrix $E\{\mathbf{R}(n)^T\mathbf{R}(n) + h\mathbf{\Delta}^T\mathbf{\Delta}\}$ is invertible, the optimum filter coefficients can be computed by substituting Eq. (27) into Eq. (25).

B. Conditions for the simulations

The simulations that follow are intended for demonstrating the benefits of the proposed technique in time domain equalization of random signals in a three-dimensional rectangular room. For this purpose, the conventional modal sum of the sound field in a lightly damped rectangular enclosure with walls of uniform specific acoustic admittance proposed by Morse[22] is used in the form presented by Bullmore *et al.*[23] The sources are modeled as square pistons that vibrate with a normal velocity $u_l = q_l/A$, where $A=a^2$ is the area of the piston sources. The piston sources are oriented inside the room so that their surfaces are parallel to the xz-plane.

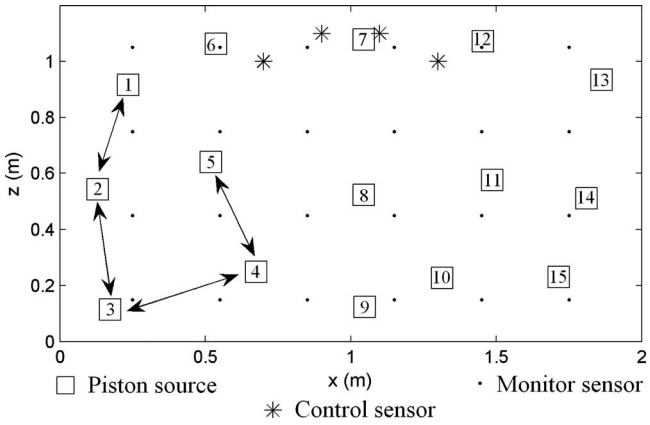

FIG. 8. Front view of the room and of the equalization system. The black arrows indicate the correlation path while moving from source 1 to source 15 on the left wall at $y=0.1$ m.

The room has dimensions of $L_x=2$ m, $L_y=3.2$ m, and $L_z=1.2$ m, and the damping ratio is set equal to 0.015, which corresponds to a reverberation time of 0.73 s at 100 Hz. The sampling rate in the simulations is 1 kHz. Equalization in an extended three-dimensional listening area inside the room is now attempted by the simulation of a plane wave traveling along the y-axis. This is to be achieved with the use of 30 piston sources, 15 on each wall perpendicular to the direction of propagation at $y=0.1$ and $y=3.1$ m. The piston sources are placed according to a 3×5 pattern on each wall, with their centers displaced 9 cm from their initial positions at a random angle on the xz-plane, as shown in Fig. 8. This random source distribution is chosen in order to avoid the ideal conditions that are met by deterministic source placement in the rectangular room. Although proper placement tactics can simplify the problem and define simple optimum correlation in the source equalization filters,[10,13] these conditions are not easily defined in the real world, for example, because of non-ideal room shape, source misplacement, and non-uniform acoustic properties of the boundaries. Any correlation that might be introduced because of the symmetries in the analytical model is here avoided by misplacement of the sources. However, care is taken for the sources not to be close to one another or close to the edges and the corners of the room.

The global reproduction error in these simulations is calculated in a three-dimensional rectangular volume of $1.5\times2.4\times0.9$ m^3. The sound pressure in this volume is sampled by a grid of $6\times8\times4=192$ monitor sensors. The monitor sensors are placed as follows: in eight sensor planes of 6×4 monitor sensors, perpendicular to the direction of propagation, with the first plane placed at $y=0.3$ m and the last one at $y=2.71$ m. Each monitor sensor plane is separated from the other by a distance of 0.343 m, which is equal to the distance that the sound travels in one sampling period. The distance between monitor sensors along the x- and z-axes is equal to 0.3 m and the monitor sensors are extended from $x=0.25$ to $x=1.75$ m and from $z=0.15$ to $z=1.05$ m. The monitor sensors are presented as black dots in Fig. 8.

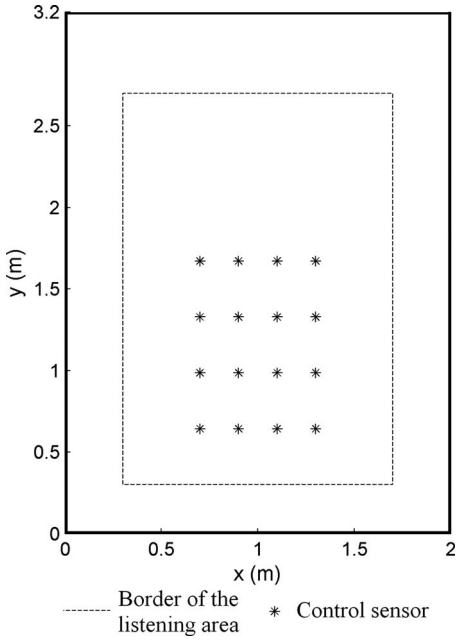

FIG. 9. Top view of the rectangular room.

Although equalization of the sound field is required in an extended three-dimensional region, this is to be achieved with the limited number of $4\times4=16$ control sensors placed, as shown in Figs. 8 and 9. The separation distance between control sensors along the y-axis is also equal to 0.343 m while the separation distance along the x-axis is equal to 0.2 m. The control sensors are located at $x=0.7$, 0.9, 1.1, and 1.3 m with the first four error sensors placed at $y=0.3$ m, and the last four ones at $y=1.329$ m (see Fig. 9). The control sensors at $x=0.7$ and 1.3 m are at a height of $z=1$ m, and those at $x=0.9$ and 1.1 m are at a height of $z=1.1$ m. The control sensors are thus placed just below the ceiling, at a position where in a more realistic case of a higher room they would not interfere with the presence and motion of listeners.

For the time domain simulations, an input signal with a delta function as the autocorrelation function is assumed, and the loudspeakers are modeled as first order analog high pass filters with a pole at 100 Hz. The frequency responses from the input of each loudspeaker to the output of each sensor are calculated with the modal model. All the natural modes up to 1100 Hz are taken into account for the simulation of the sound field in the room. These frequency responses are multiplied by the frequency response of a low pass anti-aliasing filter and by the frequency response of the loudspeakers, and an inverse fast Fourier transform (FFT) is applied for the calculation of the discrete impulse responses $c_{ml}(n)$. The number of coefficients used for each impulse response is $J=300$, and the number of coefficients in each control filter $h_l(n)$ is set equal to $I=70$. A delay of 30 samples is applied to obtain the desired signal at the first four control sensors at $y=0.3$ m, and additional delays of one, two, and three samples are used for the second, third, and fourth lines of control sensors at $y=0.643$, 0.986, and 1.329 m, respectively.

C. Implementation of effort variation regularization

A second order differential operator is constructed that counts the squared deviations of the filter samples for two individual series of 15 sources. A correlation path is specified according to the source numbering, which for the left sources 1–15 at $y=0.1$ m is shown in Fig. 8. It can be seen that sources are linked vertically rather than horizontally. A similar correlation path is applied on the other side, linking sources 16–30 at $y=3.1$ m. A 30×30 differential operator is thus constructed as

$$\mathbf{D}=\begin{bmatrix}\mathbf{D}' & \mathbf{0}\\ \mathbf{0} & \mathbf{D}'\end{bmatrix}, \tag{30}$$

where $\mathbf{D}'$ is a 15×15 tri-diagonal matrix with all elements in the main diagonal equal to -2, except elements (1,1) and (15,15), which are equal to -1, and all elements in the first two parallel diagonals are 1. The optimum equalization filters can now be calculated for effort weighting according to Eq. (26) and for the proposed technique according to Eq. (25), where the matrix $\mathbf{W}$ is constructed according to Eqs. (27) and (28), using $\mathbf{D}$ as in Eq. (30).

D. Reproduction performance

The success of each control technique is judged by the accuracy of reproduction in the entire listening area. The global reproduction error can be calculated as a function of the frequency as

$$\mathrm{ELS}(\omega)=\left(\frac{\sum_{m=1}^{M}(1-|p_m(\omega)|)^2}{M}\right)^{1/2}, \tag{31}$$

where $p_m(\omega)$ is the reproduced sound pressure at the mth monitor sensor at radial frequency ω, and $M=192$ is the number of the monitor sensors. The reproduced sound pressure $p_m(\omega)$ is determined by calculating an FFT of the signal $\hat{d}_m(n)$ detected at each sensor after equalization. This global index is intended for examining the quality of equalization achieved with the limited control sensor array previously presented. Apart from effort regularization and effort variation regularization, equalization with two individual coupled source arrays is also examined. The coupling condition assumes exactly the same equalization filter for all 15 sources at each side, specifying thus an absolute a-priori known condition which is similar to the previous example of equalization in the two-dimensional non-rectangular room. This technique has given excellent results in the case of sound equalization in a three-dimensional room with sources optimally placed on each wall.[13]

The values of the regularization parameters for effort regularization and for effort variation regularization were set equal to $\mu=10^{10}$ and $h=0.5\times10^{12}$, respectively. These values were chosen as the ones that ensure the minimum possible global reproduction error for each technique throughout the entire frequency range from 0 to 500 Hz. The variation in ELS as a function of the frequency for effort regularization, effort variation regularization, coupled source arrays, and ideal system can be seen in Fig. 10.

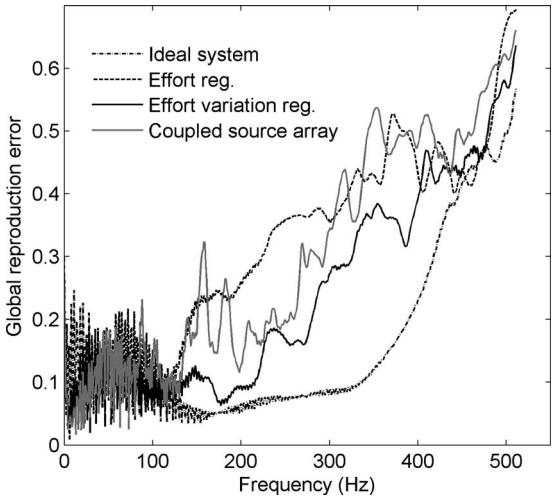

FIG. 10. Global reproduction error as a function of the frequency for effort regularization, effort variation regularization, coupled source arrays, and ideal system calculated in the time domain.

A reduction in the equalization performance with frequency is apparent with all control techniques. The performance of the ideal system indicates that good equalization is possible up to about 500 Hz, where the deviations of the reproduced sound pressure inside the listening area remain below the ±6 dB criterion. Source coupling seems to offer an advantage compared to classical effort regularization, but the global reproduction error of the proposed technique is always smaller than that of effort regularization and also that with the coupled sources. Again, this improvement is translated into reduction in equalization errors at listening positions away from the control sensors. A representative example can be seen in Fig. 11 for the monitor sensor at (1.45, 2.358, 0.45) m, illustrating the impulse and the frequency response before and after equalization for classical effort regularization (a) and (b), coupled source array (c) and (d), and effort variation regularization (e) and (f). The results

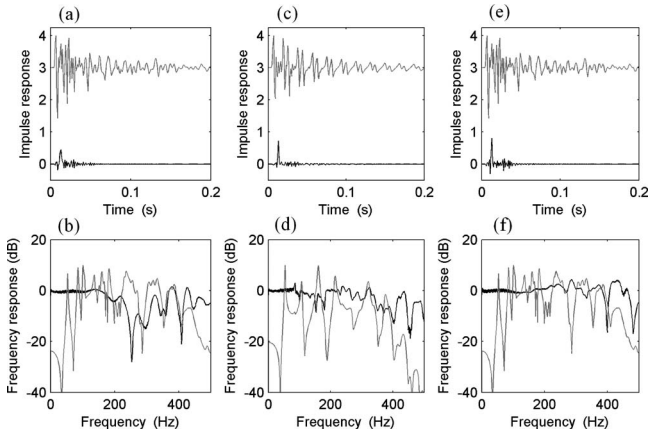

FIG. 11. Impulse and frequency responses at the monitor sensor at (1.45, 2.358, 0.45) m for effort regularization [(a) and (b)], coupled source array [(c) and (d)], and effort variation regularization [(e) and (f)]. The results before equalization in (a), (b), (e), and (f), shown with gray lines, are obtained by exciting the room with only source number 1. The results before equalization in (c) and (d) are obtained after exciting the room with all 15 left sources. The impulse responses before equalization are shown with an offset of +3 linear units for clarity.

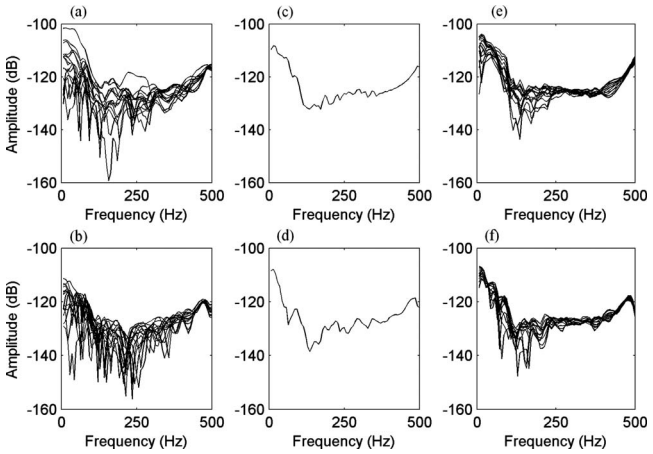

FIG. 12. Source amplitudes as functions of the frequency for effort regularization [(a) and (b)], coupled source array [(c) and (d)], and effort variation regularization [(e) and (f)]. The left sources are shown in (a), (c), and (e), and the right sources are shown in (b), (d), and (f).

before equalization, shown with gray lines, are obtained by exciting the room with piston source number "1" in (a), (b), (e), and (f) and by exciting the room with all left sources at $y=0.1$ m in (c) and (d). The impulse and frequency responses before equalization include the effect of the antialiasing filter. The impulse responses before equalization are shown with an offset of +3 linear units for the sake of clarity.

Some understanding of the effect of each control technique in the estimation of the solution $\mathbf{a}_o$ is gained by observing in Fig. 12 the variation in the amplitude (as derived from the FFT of the equalization filters impulse responses) for effort regularization in (a) and (b), coupled source array in (c) and (d), and effort variation regularization in (e) and (f). The upper row of graphs refers to the left 15 sources, and the lower one to the right sources. Although the proposed technique has reduced the amplitude variation compared to that of the classical approach in (a) and (b), a certain variation can be seen at both sides in (e) and (f) below 250 Hz, which possibly explains the avoidance of the steep error peaks observed for the coupled source case between 100 and 200 Hz in Fig. 10.

V. CONCLUSIONS

A control approach termed effort variation regularization has been proposed and examined. In the frequency domain this technique involves modifying the original cost function based on minimization of the least-squares reproduction error by adding a term that is proportional to the squared deviations between complex source strengths, calculated independently for the sources at the two walls perpendicular to the direction of propagation. Simulation results in a two-dimensional non-rectangular room have shown that effort variation regularization results in equalization of the phase of the complex strengths on each side of the room, without preventing the amplitude deviations that are necessary for room compensation. Applied to the time domain equalization of broadband signals, effort variation regularization has been successfully realized by independently penalizing the squared deviations of the source equalization filter coeffi-

cients at each side of the room. The proposed technique has been demonstrated to lead to smaller global reproduction error and thus better equalization performance at listening positions away from the control region than effort regularization and a simple coupling source array method.

[1]J. Mourjopoulos, P. M. Clarkson, and J. K. Hammond, "A comparative study of least-squares and homomorphic techniques for the inversion of mixed phase signals," in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP' 82), Paris, France (1982), Vol. **7**, pp. 1858–1861.

[2]J. Mourjopoulos, "On the variation and invertibility of room impulse response functions," J. Sound Vib. **102**, 217–228 (1985).

[3]A. Mäkivirta, P. Antsalo, M. Karjalainen, and V. Valimaki, "Modal equalization of loudspeaker-room responses at low frequencies," J. Audio Eng. Soc. **51**, 324–343 (2003).

[4]M. Karjalainen, P. Atsalo, and A. Mäkivirta, "Modal equalization by temporal shaping of room responses," in Proceedings of the AES 23rd International Conference, Copenhagen (2003).

[5]J. Wilson, M. Capp, and R. Stuart, "The loudspeaker-room interface-controlling excitation of room modes," in Proceeding of the AES 23rd International Conference, Copenhagen (2003).

[6]S. J. Elliott and P. A. Nelson, "Multiple-point equalization in a room using adaptive digital filters," J. Audio Eng. Soc. **37**, 899–907 (1989).

[7]F. Asano and D. Swanson, "Sound equalization in enclosures using modal reconstruction," J. Acoust. Soc. Am. **98**, 2062–2069 (1995).

[8]A. O. Santillán, "Spatially extended sound equalization in rectangular rooms," J. Acoust. Soc. Am. **110**, 1989–1997 (2001).

[9]J. C. Sarris, F. Jacobsen, and G. E. Cambourakis, "Sound equalization in a large region of a rectangular enclosure," J. Acoust. Soc. Am. **116**, 3271–3274 (2004).

[10]A. O. Santillán, C. S. Pedersen, and M. Lydolf, "Experimental verification of a low-frequency global sound equalization system based on free field propagation," Appl. Acoust. **68**, 1063–1085 (2007).

[11]A. Celestinos and S. Birkedal Nielsen, "Controlled acoustic bass system (CABS), a method to achieve uniform sound field distribution at low frequencies in rectangular rooms," J. Audio Eng. Soc. **56**, 915–931 (2008).

[12]N. Stefanakis, J. Sarris, G. Cambourakis, and F. Jacobsen, "Power-output regularization in global sound equalization," J. Acoust. Soc. Am. **123**, 33–36 (2008).

[13]N. Stefanakis, J. Sarris, and G. Cambourakis, "Source placement for equalization in small enclosures," J. Audio Eng. Soc. **56**, 357–371 (2008).

[14]P. A. Nelson, "A review of some inverse problems in acoustics," Int. J. Acoust. Vib. **6**, 118–134 (2001).

[15]S. J. Elliott, *Signal Processing for Active Control* (Academic, London, 2001).

[16]P. A. Gauthier, A. Berry, and W. Woszczyk, "Sound-field reproduction in-room using optimal control techniques: Simulations in the frequency domain," J. Acoust. Soc. Am. **117**, 662–678 (2005).

[17]P. C. Hansen, "Perturbation bounds for discrete Tikhonov regularization," Inverse Probl. **5**, L41–L44 (1989).

[18]M. R. Bai, "Study of acoustic resonance in enclosures using eigenanalysis based on boundary element methods," J. Acoust. Soc. Am. **91**, 2529–2538 (1992).

[19]N. Stefanakis and J. Sarris, "Sound field reproduction using the boundary element method," in Proceedings of the ICSV13-Vienna (2006).

[20]D. L. Kepler and D. W. Steele, "Constant directional characteristics from a line source array," J. Audio Eng. Soc. **11**, 198–202 (1963).

[21]P. A. Nelson, F. Orduña-Bustamante, and H. Hamada, "Multichannel signal processing techniques in the reproduction of sound," J. Audio Eng. Soc. **44**, 973–989 (1996).

[22]P. M. Morse, *Vibration and Sound*, 2nd ed. (McGraw-Hill, New York, 1948).

[23]A. J. Bullmore, P. A. Nelson, A. R. D. Curtis, and S. J. Elliott, "The active minimization of harmonic enclosed sound fields, part II: A computer simulation," J. Sound Vib. **117**, 15–33 (1987).